<u>IN THE CLAIMS</u>

Please amend the claims as follows:

Claim 1 (Currently Amended): An identifier generating method comprising:

a canonicalization process step of subjecting document data to a canonicalization

process to correct fluctuation of expression <u>including at least one of deletion of redundant</u>

<u>white spaces and recovery from omission of a close tag</u>; and

an identifier generating step of generating an identifier uniquely specifying the

document data or part thereof, based on all or part of the document data having been

subjected to the canonicalization process in the canonicalization process step.

Claim 2 (Currently Amended): An identity determining method comprising:

a canonicalization process step of subjecting document data to a canonicalization

process to correct fluctuation of expression <u>including at least one of deletion of redundant</u>

<u>white spaces and recovery from omission of a close tag</u>;

an identifier generating step of generating an identifier uniquely specifying the

document data or part thereof, based on all or part of the document data having been

subjected to the canonicalization process in the canonicalization process step; and

an identity determining step of determining whether there is a common portion

between one document data and another document data, based on the identifier having been

generated in the identifier generating step.

Claim 3 (Original): An identity determining method comprising:

an identifier generating step of, based on all or part of encoded data of document data,

generating an identifier uniquely specifying the document data or part thereof; and

an identity determining step of determining whether there exists a common portion between one document data and another document data, based on the identifier having been generated in the identifier generating step.

Claim 4 (Currently Amended): An identifier transmitting method comprising:

a canonicalization process step of subjecting document data to a canonicalization process to correct fluctuation of expression including at least one of deletion of redundant white spaces and recovery from omission of a close tag;

an identifier generating step of generating an identifier uniquely specifying the document data or part thereof, based on all or part of the document data having been subjected to the canonicalization process in the canonicalization process step; and

an identifier transmitting step of transmitting the identifier having been generated in the identifier generating step.

Claim 5 (Currently Amended): An identifier generating apparatus comprising:

canonicalization process means for subjecting document data to a canonicalization process to correct fluctuation of expression including at least one of deletion of redundant white spaces and recovery from omission of a close tag; and

identifier generating means for generating an identifier uniquely specifying the document data or part thereof, based on all or part of the document data having been subjected to the canonicalization process by the canonicalization process means.

Claim 6 (Original): The identifier generating apparatus according to Claim 5, wherein the canonicalization process means comprises type standardization process means for, using a class definition file of the document data describing a definition of a type,

standardizing a type of expression for a value described in the document data, in accordance with the type defined by the class definition file.

Claim 7 (Original): The identifier generating apparatus according to Claim 6, wherein the type standardization process means is configured to standardize an accuracy of numerical data described in the document data, in accordance with a definition of a type for numerical data described in the class definition file of the document data.

Claim 8 (Original): The identifier generating apparatus according to Claim 5, wherein the canonicalization process means comprises document data generating means for transforming first partial data and second partial data into document data in accordance with a predetermined transformation rule.

Claim 9 (Currently Amended): An identity determining apparatus comprising:

canonicalization process means for subjecting document data to a canonicalization process to correct fluctuation of expression including at least one of deletion of redundant white spaces and recovery from omission of a close tag;

identifier generating means for generating an identifier uniquely specifying the document data or part thereof, based on all or part of the document data having been subjected to the canonicalization process by the canonicalization process means; and

identity determining means for determining whether there is a common portion between one document data and another document data, based on the identifier having been generated by the identifier generating means.

Claim 10 (Original): The identity determining apparatus according to Claim 9, wherein the canonicalization process means comprises type standardization process means for, using a class definition file of the document data describing a definition of a type, standardizing a type of expression for a value described in the document data, in accordance with the type defined by the class definition file.

Claim 11 (Original): The identity determining apparatus according to Claim 10, wherein the type standardization process means is configured to standardize an accuracy of numerical data described in the document data, in accordance with a definition of a type for numerical data described in the class definition file of the document data.

Claim 12 (Original): The identity determining apparatus according to Claim 9, wherein the canonicalization process means comprises document data generating means for transforming first partial data and second partial data into document data in accordance with a predetermined transformation rule.

Claim 13 (Original): The identity determining apparatus according to Claim 9, further comprising a cache preliminarily storing the identifier having been generated by the identifier generating means, in correlation with the document data or a result of a predetermined process on the document data,

wherein the identity determining means is configured to:

perform a search inside the cache on the basis of the identifier of the document data as a target for a determination on identity, which has been generated by the identifier generating means, determine that there exists identical document data, if the same identifier as the

aforementioned identifier is present, and determine that there exists no identical document

data, if the same identifier as the aforementioned identifier is absent.

Claim 14 (Original): The identity determining apparatus according to Claim 9,

further comprising second identifier generating means for generating an identifier uniquely

specifying the document data or part thereof, based on all or part of the document data, prior

to execution of the canonicalization process by the canonicalization process means,

wherein identity between one document data and another document data is

determined on the basis of the identifier having been generated by the second identifier

generating means and wherein if they are determined not to be identical, the canonicalization

process means performs the canonicalization process of the document data.

Claim 15 (Currently Amended): An identity determining apparatus comprising:

canonicalization process means for subjecting document data to a canonicalization

process to correct fluctuation of expression including at least one of deletion of redundant

white spaces and recovery from omission of a close tag;

identifier generating means for, based on all or part of encoded data of document data,

generating an identifier uniquely specifying the document data or part thereof; and

identity determining means for determining whether there exists a common portion

between one document data and another document data, based on the identifier having been

generated by the identifier generating means.

Claim 16 (Original): The identity determining apparatus according to Claim 15,

wherein an instruction to skip a process of decoding the encoded data of the document data is

issued when the identity determining means determines that the two document data are

identical.

Claim 17 (Currently Amended): An identifier transmitting apparatus comprising:

canonicalization process means for subjecting document data to a canonicalization

process to correct fluctuation of expression including at least one of deletion of redundant

white spaces and recovery from omission of a close tag;

identifier generating means for generating an identifier uniquely specifying the

document data or part thereof, based on all or part of the document data having been

subjected to the canonicalization process by the canonicalization process means; and

identifier transmitting means for transmitting the identifier having been generated by

the identifier generating means.

Claim 18 (Original): The identifier transmitting apparatus according to Claim 17,

wherein the identifier transmitting means comprises transmitted data generating means for

generating data obtained by replacing all or part of the document data by the identifier.

Claim 19 (Original): The identifier transmitting apparatus according to Claim 18,

wherein the transmitted data generating means is configured to generate transmitted data

described by an identifier uniquely specifying partial data included in the document data, and

difference data between the partial data and the document data.